DOCUMENT RESUME

ED 059 547 24 EA 004 091

AUTHOR	Hanson, Ralph A.; Berger, Robert J.
TITLE	Quality Assurance in Large Scale Installations of Criterion Referenced Instructional Programs.
INSTITUTION	Southwest Regional Educational Lab., Inglewood, Calif.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO	SWRL-PP-18
BUREAU NO	BR-6-2865
PUB DATE	Aug 71
CONTRACT	OEC-4-7-062865-3073
NOTE	12p; Paper presented at National Council on Measurement in Education annual meeting (New York City, February 5-7, 1971)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Criterion Referenced Tests; Data Analysis; *Educational Quality; *Instructional Technology; *Performance Criteria; *Quality Control; Sampling

ABSTRACT

Quality assurance is viewed as a means of maintaining desired performance levels during the operational use of a developed instructional program. Presented here is an analysis of the functions and the basic requirements of a quality assurance system. Procedures for implementing the system are described and discussed. The procedures include specification and development of performance indicators, sampling, data collection and analysis, and decision rules and specifications of actions to be taken. (Author)

PERMISSION TO REPRODUCE THIS COPY RIGHTED MATERIAL HAS BEEN GRANTED BY Southwest Regional Laboratory TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

© 1971 SOUTHWEST REGIONAL LABORATORY
Published by Southwest Regional Laboratory, 11300 La Cienega Blvd., Inglewood, California 90304.
Printed by Educational Technology Publications, Englewood Cliffs, New Jersey 07632.
All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

BR 6-2865
PA 24

ED 059547

SOUTHWEST REGIONAL LABORATORY
Professional Paper 18
August 1971

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

QUALITY ASSURANCE IN LARGE SCALE INSTALLATIONS OF CRITERION REFERENCED INSTRUCTIONAL PROGRAMS

Ralph A. Hanson and Robert J. Berger

ABSTRACT

Quality Assurance is viewed as a means for maintaining desired performance levels during the operational use of a developed instructional program. Presented here is an analysis of the functions and the basic requirements of a Quality Assurance system. Procedures for implementing the system are described and discussed. The procedures include specification and development of performance indicators, sampling, data collection and analysis, and decision rules and specifications of actions to be taken.

EA 004 091

QUALITY ASSURANCE IN LARGE SCALE INSTALLATIONS OF CRITERION REFERENCED INSTRUCTIONAL PROGRAMS[1]

Ralph A. Hanson and Robert J. Berger

The two major purposes of this paper are to analyze the nature and function of Quality Assurance procedures in instructional technology, and to describe the strategy used to develop a simple Quality Assurance system for an instructional program.

The Quality Assurance system described here is currently being used with the Southwest Regional Laboratory's First Year Communication Skills Program, a criterion referenced instructional program (instructional programs incorporating prespecified performance criteria and procedures for periodic assessment of the attainment of instructional outcomes) used by over 30,000 pupils in 59 school districts. The process of placing this and other criterion referenced instructional programs into the school environment is referred to as installation. Product (or program) installation is successfully accomplished when the instructional program performs at or above the prespecified criterion levels without direct control or influence by the developers of the program.

Quality Assurance itself typically comprises the set of procedures used by the schools to assess the reliability with which instructional outcomes are being attained during the operational use of the program. Specification of data to be collected, decision rules to be applied, and actions to be taken to avoid program failures is inherent in any Quality Assurance system.

Quality Assurance Implementation Procedures

The complete set of procedures involved in implementing a Quality Assurance system includes seven major components.

Specification of Indicator Variables

A number of different aspects of an installed instructional product exist which determine how well it performs. Indicator variables measuring the important attributes of each of these different aspects should be defined. If, for example, personnel training is a requirement

---

[1]Presented at the annual meeting of the National Council on Measurement in Education held in New York, New York, February 5-7, 1971.

of successful program performance, then variables indicating how well this training has been carried out must be defined. The indicator variables in this case can be paper and pencil tests of the training program content, observational measures of the trained personnel, or other such measures. The specific variables included should be those which provide the most valid information on each program-critical aspect, such as:

1. Pacing: measures of instructional time in hours, days, weeks, etc.

2. Performance: en route measures of learning, such as unit or midterm tests, observations, etc.

3. Logistics: indicator reports of failure to deliver materials, errors in materials, etc.

Measurement of Indicator Variables

The measurement units chosen for indicator variables will necessarily vary since some are performance measures and others are status indicators. In all cases, measures should be developed so that a criterion or acceptability threshold can be established.

Definition of Decisions Rules

While one can define several threshold acceptability levels associated with different degrees of difficulty, the most basic are those which signal a major program failure. A major program failure is defined differentially depending on the specific objectives of a program. In most cases, however, the primary criteria for a major program failure are determined by the proportion of schools or classes in which the end of program pupil performance does not reach the prespecified criterion. Variable levels associated with program failure in earlier developmental tryouts are usually excellent indicators of potential failure in subsequent use of the program.

Variable cutoff points can also be established on an analytical basis. Suppose, for example, a program is comprised of ten instructional units of equal length, to be completed in a fixed period of time. Under these conditions, one can quite easily determine cutoff points for pacing which meet any deadline and ensure program completion.

Cutoff points are established for such simple, yet critical events as a delay in the receipt of instructional materials. It is clear that if the materials do not arrive on time, the program will have a late start and it will have to be accelerated to assure completion. In such

cases cutoff points are made without any prior information. As more information becomes available about the performance characteristics of a program and its relationship to each indicator variable, adjustments in such cutoff points are made.

Sampling Procedures

Several sampling considerations are important in a Quality Assurance system. One consideration is the choice of the number of program participants actually submitting data. With large scale use of a program, the component of data available is usually far greater than the amount required to provide accurate estimates of the indicator variables. Thus, only a fraction of the total participant group may be required to provide Quality Assurance data.

A second consideration is that different sampling units may be appropriate for the different indicator variables being estimated. If, for example, the indicator variable is a measure of whether program materials were delivered to districts, one would want the accuracy to be extremely high. The failure of one district to receive program materials inevitably leads to the omission of results for thousands of students. Therefore, a large percentage, if not all of the districts using the program, should be sampled in this case.

Another consideration in sampling is the amount of information necessary from each participant. If the amount and time required to supply the information is high, it may be better to have a number of users supply information on different variables rather than one user supplying information on all variables. Other sampling considerations can be identified in the context of a specific Quality Assurance system.

Collecting Quality Assurance Data

Data collection for Quality Assurance has certain special problems. Program users often have limited commitments to quality standards viewed as important by the developer. The data collected must be processed and reported within a short period of time in order to be useful for Quality Assurance. Data must be gathered as unobtrusively as possible in order to present a valid picture of program status.

In designing a system to satisfy these conditions, a number of principles are applicable.

1. Minimize the burden on each participant by collecting only the data required.

2. Use forms and data collection procedures which can be easily and quickly fulfilled by users.

3. Include, when possible, indicator variables which can be gathered routinely by participants as part of their regular program or school procedures. Data which require a special effort by participants are least likely to be completed and may alter use of the program.

Analysis and Summarization of Data

Data analysis requirements for Quality Assurance can be quite stringent. Data on some indicator variables must be analyzed immediately upon receipt in order to be of value for decision making. Meeting this requirement may necessitate weekly summaries generated and updated on a daily basis. If large volumes of data are involved, it may be most efficient to have both the analysis and flagging of important results carried out by computer.

Specification of Actions to be Taken

The final step in the Quality Assurance implementation process is the specification of the appropriate actions to be taken in the case of a major program failure. A number of options are available depending on the severity and generalizability of the failure.

If it is found that the program failure is generalizable across all users, several alternatives are available depending on the severity of the failure. A relatively minor program failure requires only a letter to users suggesting a procedural change. On the other hand, if the program failure is both generalizable and severe, the response requires dispatching personnel for further troubleshooting. Such data indicate the need for inservice training or other revisions to the program.

If the program failure, whether minor or major, is limited to a small proportion or subgroup of program users, intervention procedures need be directed only to those users.

Information from Quality Assurance interventions should be fed back into the development cycle aiding in program revisions.

Development of a Simple Quality Assurance System

The procedures just described have been followed in developing a simple Quality Assurance System for the Southwest Regional Laboratory's First Year Communication Skills Program. This program is designed to teach kindergarten children four basic reading skills:

1. A reading vocabulary of approximately 100 words.

2. The ability to read 23 selected initial and ending sounds.

3. The ability to sound out and read any one-syllable word composed of word elements from the program.

4. The ability to name each letter of the alphabet, when shown the printed letter.

Program instruction is divided into ten units, each ending with a criterion referenced test covering all program objectives. The complete instructional system for the First Year Communication Skills Program includes several subsystems, including those concerned with teacher and supervisor training, supervisor management, and instructional support. In this first application of Quality Assurance procedures, however, not all of these systems were considered separately. Instead, a simple system was designed which gathers data on two basic aspects of the program: the completion of the ten instructional units (pacing), and overall student mastery of program objectives. The major objectives of this system are to ensure that:

1. Instructional proficiency on the four outcomes of each unit is attained by each pupil.

2. All ten units of the program are completed during the school year.

Specification of Indicator Variables

Student Mastery is measured by performance on the unit criterion tests and the pacing objective is measured by the rate of completion of program units.

Development of Measures of Each Indicator Variable

Performance on Unit Criterion Tests is measured simply by recording the number of items answered correctly on each of the four test outcomes. Since there are five item outcomes, scores range from 0 to 5 correct. Completion of program units is measured by recording days spent per unit. A unit is considered to be complete when the last set of students (who are not makeups or absentees) have received the criterion test for that unit.

Decision Rules for Indicator Variables

Performance on unit criterion tests. In order to derive an acceptability threshold for the unit tests, an analysis of data from the 1969-70 program tryout was carried out. This tryout provided information on both the unit tests and on an end of year final test for a sample of classes.

The analysis considered only classes which had completed all ten units of the program and posed the following question: On how many units can a class fail to reach an 80% mastery level and still have a chance of reaching criterion on the end of year test?

The procedure followed was to look at classes which had failed to reach criterion on differing numbers of units, noting the proportion in each category above the acceptability threshold on the end of year test. The results of this analysis are given in Table 1.

As this table indicates, a class which fails to reach the criterion on more than two of the unit tests has little chance of mastering the end of program test. Thus, the cutoff point for this variable indicator was placed at 2. Classes which fall below the criterion on more than two units are to be considered as performing below the criterion in terms of pupil mastery.

Completion of program units. In order to determine empirically the cutoff point for completion of program units an analysis was carried out of data collected in the 1969-70 tryout of the FYCSP program. This analysis included data from 10 classes which had completed all 10 units of the program. The objective of the analysis was to specify the minimum acceptable completion data for each program unit.

The analyses specified the minimum number of days required to complete all remaining units after completion of a unit. It was obtained by counting the number of days required by the fastest class in last years tryout finishing the remaining units. This figure will be used to determine if each Quality Assurance class can still complete the program after each unit completed. These minimum number of days are given in Table 2 based on the data illustrated in Figure 1.

The procedure to be followed is to determine the number of school days left for a given class after the completion of a unit of the program. If this number is more than the number required by the fastest class using the program last year (minumum number of days) then the class will be considered on schedule. If it falls below this minimum it will be flagged and designated behind schedule in terms of program completion.

Specification of the Sample

In specifying the Quality Assurance for the 1970-71 year, several decisions were made at the outset. The first was that the school should be the unit in collecting data. This was decided because program options and training are distributed to schools rather than classes or districts. The second decision was to include 10 percent of the participating

Table 1

Number of Classes Mastering the Posttest as Related to the Number of Unit Tests Mastered

Number of Unit Tests Mastered	No. of Classes	No. Mastering Posttest	Percentage Mastering the Posttest
5	4	1	25
6 or 7	2	1	50
8 or 9	4	4	100
10 Units	9	9	100
Total	19	15	

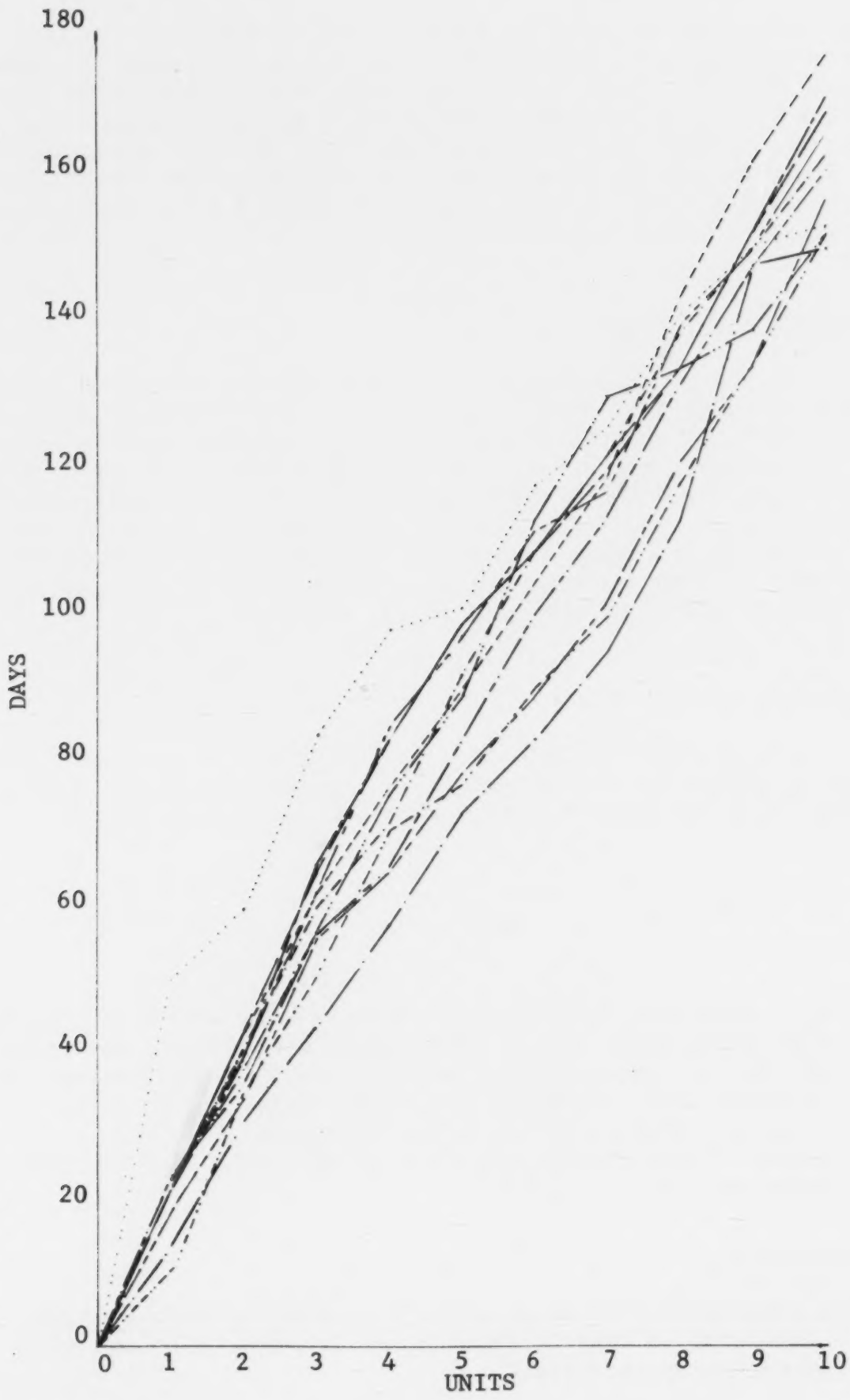
Table 2

Completion of Program Units

Number of Unit Completed	Minimum Number of Days Required to Complete the Program by Fastest Class from 1969-70*
0	151
1	119
2	108
3	93
4	79
5	68
6	57
7	38
8	18
9	4
10	0

*Based on results from 10 classes which completed all units.

Fig. 1. Number of days required to complete each FYCSP unit.



schools in Quality Assurance. Since there are over 300 schools participating, this made the Quality Assurance sample about 30 schools. Any larger sample would make the analysis requirements very extensive and any less would not be adequate to sample the universe.[2]

The procedure followed in sampling was to represent as many districts as possible. Specifically, an attempt was made to represent all large districts (more than 10 schools), some medium sized districts (5 to 10 schools) and a sample of the small districts (less than 5 schools). Further, the schools chosen from a particular district were compared to the district average on the basis of performance on prior standardized reading tests. Those drawn were changed to be more representative if the data indicated the need.

Data Gathering Procedures

The procedure to be followed in gathering the data from Quality Assurance schools is to have teachers submit the Criterion Exercises for their class at the end of each unit. If classes include groups, moving at different rates, data are submitted after each group finishes. Along with the Criterion Tests themselves, teachers indicate the date of administration and if the tests are for the last group for that unit. Using this procedure, information on both unit test performance and the data of completion of the unit is obtained at the same time with a minimum effort on the part of the schools.

Summarization of the Data

The data for Quality Assurance is summarized in a monthly report. This report indicates the number of classes and schools by district which are below the minimum performance and pacing acceptability threshold.

Discussion

Quality Assurance applied to the operational use of criterion reference programs would appear to fulfill several important functions. It ensures that the specified performance level of the instructional program is maintained throughout its useful life. It provides a mechanism for systematically improving a program after its operational use has begun. Thus, changes may be made and their effects monitored in new installations.

---

[2]The results of this first attempt at Quality Assurance may indicate that some of these decisions are in error and need to be altered in future applications.